

100

104

YEUNG AND YEO: VIDEO VISUALIZATION

nance generally capture well the relative importance of video content in a wide variety of video materials."

In conclusion, from the analysis, characterization, and modeling of video, we can build intuitive and meaningful representations of video. Together they contribute to semantic visualization of video and succinct presentation of the pictorial content, as well as provide the tools for many applications in digital video, including nonlinear video access, efficient storage and retrieval, effective query, browsing and navigation, compaction, data management, and organization. These tools are the enabling technologies for new applications and services of digital video.

ACKNOWLEDGMENT

The authors thank B. Liu for his careful reading of M. M. Yeung's thesis, from which part of the manuscript is adapted, Dr. J. L. Koulouberis of IBM T. J. Watson Research Center for his help in extracting MPEG-1 video sequences from CDDROM's, guest editor Dr. H.-D. Lin for his efforts in coordinating the review, and the reviewers for their comments and suggestions.

REFERENCES

- [1] C. L. Compton and P. D. Bosc, "Internet CNN NEWSROOM: A digital video news magazine and library," in *Int. Conf. Multimedia Computing and Systems*, 1996, pp. 296-301.
- [2] B. Shabirany and D. Gibbo, "Automatic generation of pictorial transcripts of video programs," in *Multimedia Computing and Networking 1995*, vol. SPIE-2417, Feb. 1995, pp. 512-528.
- [3] P. England, R. B. Allen, M. Sullivan, and A. Heybey, "iBrowse: The Bellcore video library toolkit," in *Proc. SPIE Storage and Retrieval for Still Image and Video Databases*, Feb. 1995, vol. SPIE-2670, pp. 254-264.
- [4] M. Miller, J. Cohen, and Y. Y. Wong, "A transfer tool for video data," in *Proc. ACM Computer Graphics Workshop (CHI)*, May 1992, pp. 93-98.
- [5] H. Ueda, T. Miyazaki, and T. Yoshizawa, "IMPACT: An interactive multimedia picture dedicated multimedia authoring system," in *Proc. ACM CHI'91*, 1991, pp. 343-350.
- [6] H. Ueda, T. Miyazaki, S. Sumino, and A. Nagasaki, "Automatic structure visualization for video editing," in *Proc. INTERCHI'93*, ACM, pp. 137-141.
- [7] S. W. Seo and H. J. Zhang, "Content-based video indexing and retrieval," *IEEE Multimedia*, pp. 62-72, Summer 1994.
- [8] F. Amara, R. Deponnier, A. Hsu, and M. Y. Chu, "Content-based browsing of video sequences," *ACM Multimedia 94*, pp. 97-103, Aug. 1994.
- [9] B. Fluck and K. Karmouch, "A multimedia news delivery system over an ATM network," in *Int. Conf. Multimedia Computing and Systems*, 1995, pp. 56-63.
- [10] L. Teodoro and W. Bender, "Salient video stills: Content and context preserved," in *Proc. Multimedia 93*, pp. 39-46.
- [11] Y. Tatemura, A. Akana, K. Oishi, and T. Sadaoka, "VideoMAP and VideoSpec: Tools for automating video content," in *INTERCHI'93, Conf. Human Factors in Computing Systems*, pp. 131-136.
- [12] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, H. Qian, B. Doni, M. Cockcroft, J. Hafner, D. Lee, D. Fodorovic, D. Steele, and P. Yanke, "Query by image and video content: The QBC system," *Comput.*, vol. 28, pp. 23-32, Sept. 1993.
- [13] M. A. Smith and T. Kanade, "Video skimming for quick browsing based on audio and image characterization," *Cornell Mellon Univ., Tech. Rep. CMU-CS-95-146*, July 1995.

"Capturing 'true' semantic significance in video automatically is a difficult problem that warrants further research."

- [14] M. M. Yeung, B. L. Yeo, W. Yip, and B. Liu, "Video browsing using clustering and scene transitions on compressed sequences," in *Multimedia Computing and Networking 1995*, vol. SPIE-2417, Feb. 1995, pp. 399-413.
- [15] M. M. Yeung and B. L. Yeo, "Time-constrained clustering for segmentation of video into story units," in *Int. Conf. Pattern Recognition (ICPR'96)*, Aug. 1996, vol. C, pp. 373-380.
- [16] B. L. Yeo and B. Liu, "Rapid scene analysis on compressed videos," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, pp. 533-544, Dec. 1995.
- [17] M. M. Yeung and B. L. Yeo, "Efficient searching and clustering of video shots," in *Int. Conf. Image Processing*, vol. 1, 1995, pp. 338-341.
- [18] M. M. Yeung and B. L. Yeo, "Video content characterization and compaction for digital library applications," in *Proc. SPIE Storage and Retrieval for Still Image and Video Databases V*, vol. SPIE-3022, Feb. 1997, pp. 45-53.
- [19] H. I. Fineberg and B. D. Inala, *Visual Editing—A Graphic Guide for Journalists*. Belmont, CA: Wadsworth, 1990.
- [20] G. K. Wallace, "The JPEG still picture compression standard," *Commun. ACM*, vol. 34, pp. 30-44, Apr. 1991.



Minerva M. Yeung (S'90-M'96) received the Ph.D. and Master of Arts degrees from the Department of Electrical Engineering, Princeton University, Princeton, NJ, in 1996 and 1994, respectively, and the Bachelor of Science degree in electrical engineering (with Highest Distinction) from Purdue University, West Lafayette, IN, in 1992.

She is currently a Research Staff Member with IBM Thomas J. Watson Research Center, Yorktown Heights, NY. She worked at Image Library Applications group of IBM Research (June 1993 to September 1995), and at the Imaging Department of Siemens Corporate Research, Princeton, NJ (May 1993 to August 1993). Her research interests are in the areas of image/video processing, multimedia databases, video browsing and retrieval, multimedia data security, digital communications, and issues related to the digital libraries.

Dr. Yeung was an Intel Foundation Graduate Fellow (1995 to 1996), a Sony Graduate Fellow (1994 to 1995), and a C. W. Chu Foundation Scholar (1988 to 1992).



Bone-Lock Yeo (S'91-M'96) received the B.S.E.E. degree in electrical engineering (with highest distinction) from Purdue University, West Lafayette, IN, in August 1992, and the M.A. and Ph.D. degrees from Princeton University, Princeton, NJ, in June 1994 and January 1996, respectively.

He served in the Singapore Armed Forces from December 1986 to June 1989. He worked at C-Cube Microsystems, San Jose, CA, in Summer 1990, at Siemens Corporate Research, Princeton, NJ, in Summer 1993, and at Mathematical Sciences Research Center, AT&T Bell Laboratories, Murray Hill, NJ, June-December 1995. In December 1995, he joined IBM T. J. Watson Research Center, Hawthorne, NY, as a Research Staff Member. He is now the Manager of the Visual Computing and Communications department. His research interests include signal and image processing, data compression and communications, visualization, computer vision, and problems related to multimedia information systems.

He is serving as a Guest Editor of a special issue of *Computer Vision and Image Understanding on Computer Vision Applications for Network-Centric Computing*. He was the recipient of the 1995 IEEE Circuits and Systems Society Video Technology Transactions Best Paper Award, was the recipient of the Wallace Memorial Fellowship in Engineering (1995 to 1996), the IBM Graduate Fellowship (1994 to 1995), and was a Singapore Technologies Overseas Scholar (1989 to 1991).

Fig. 1

200

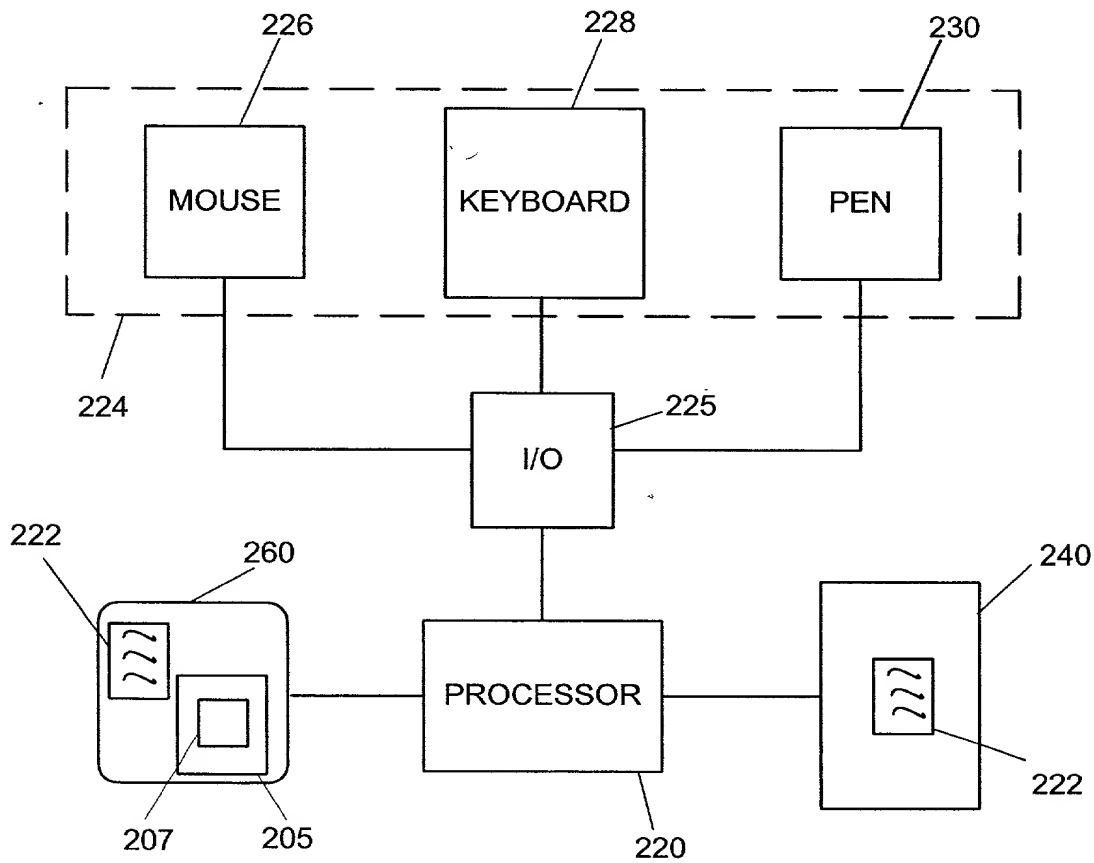


FIG. 2

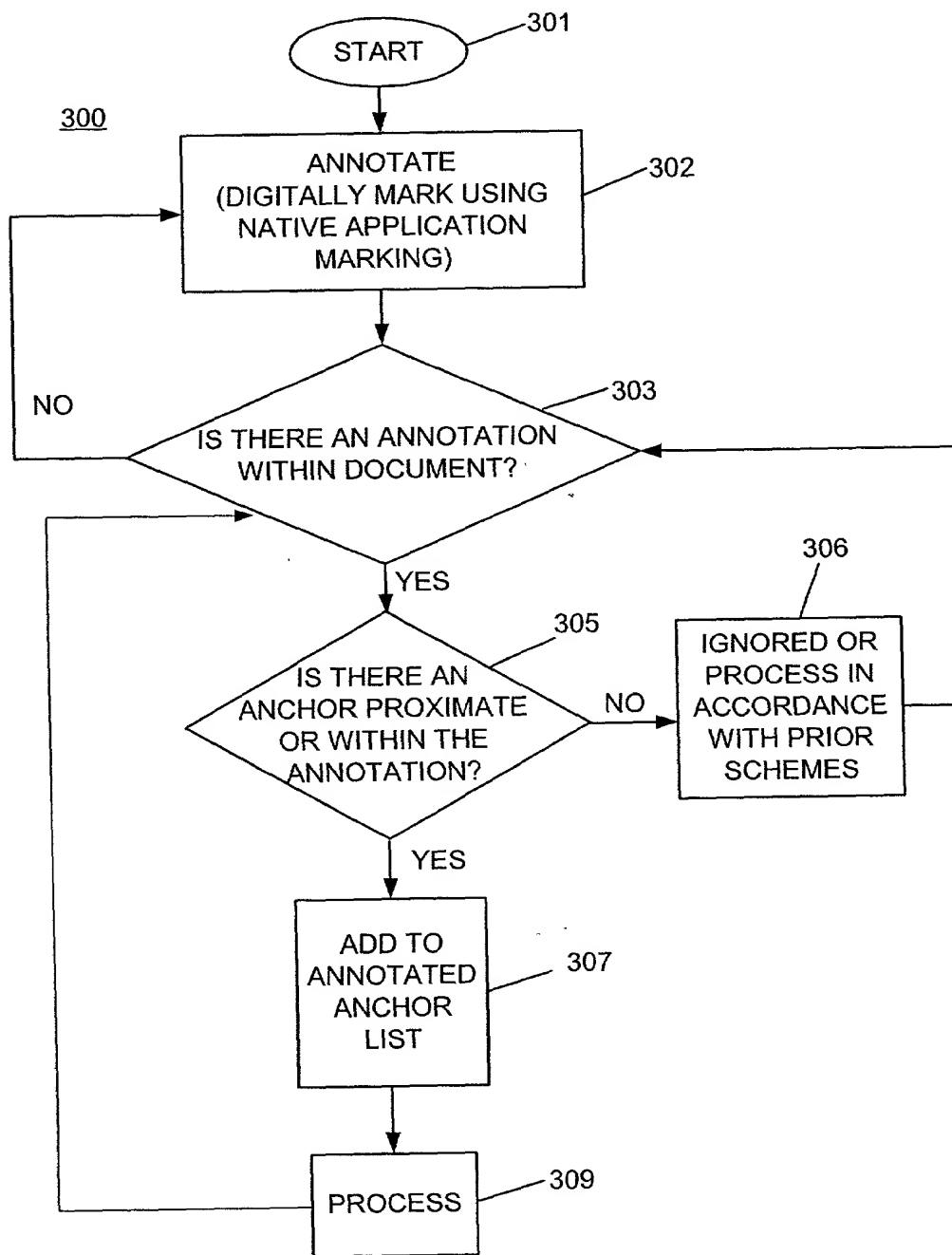


FIG. 3

400

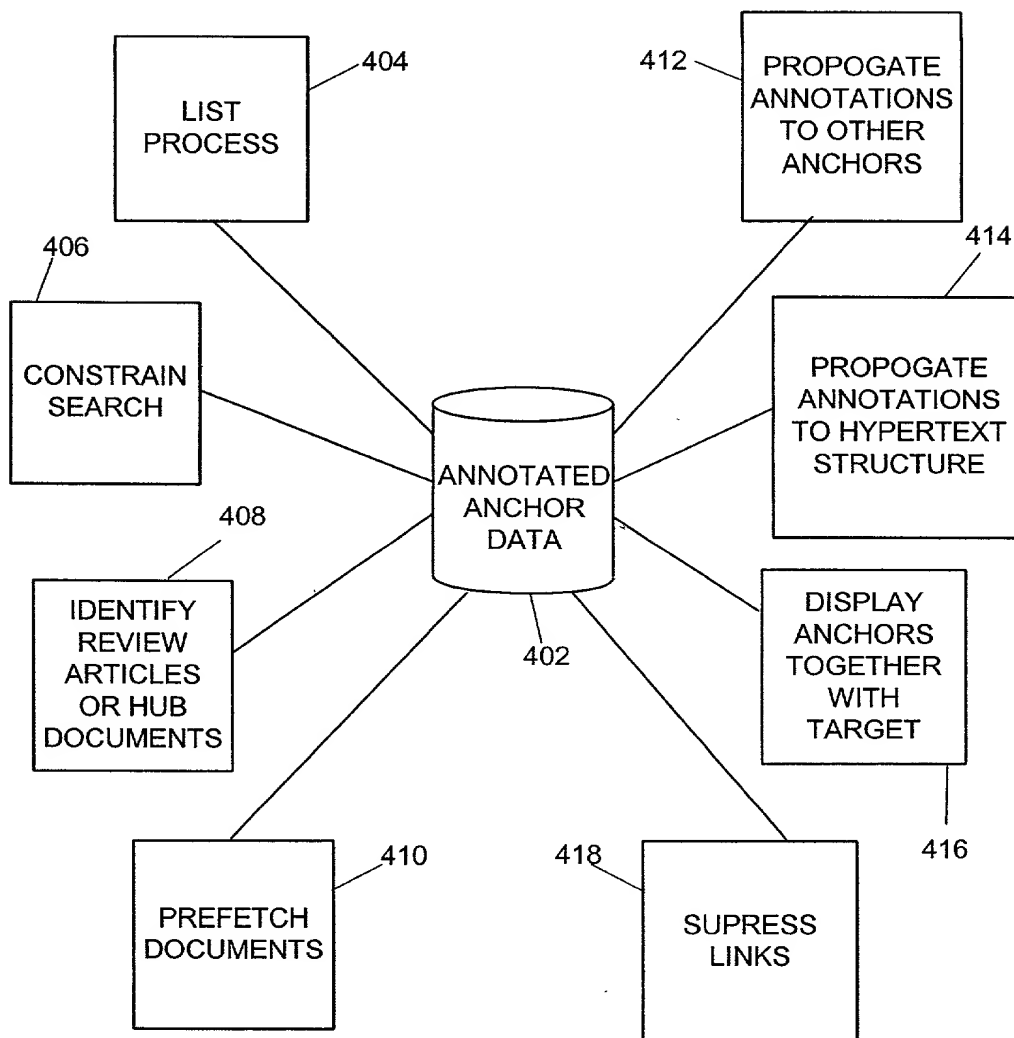


FIG. 4

404

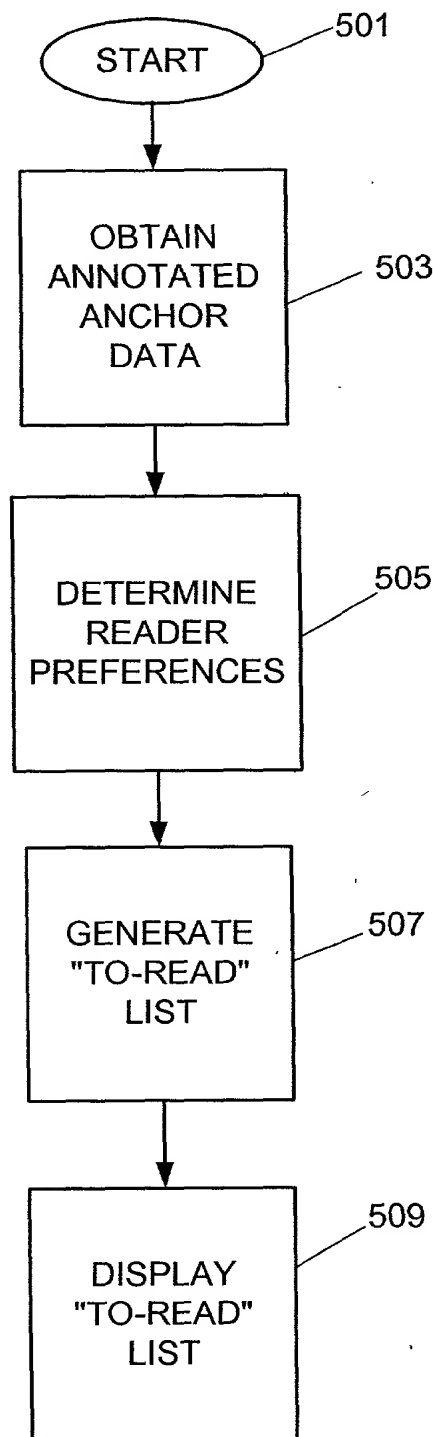
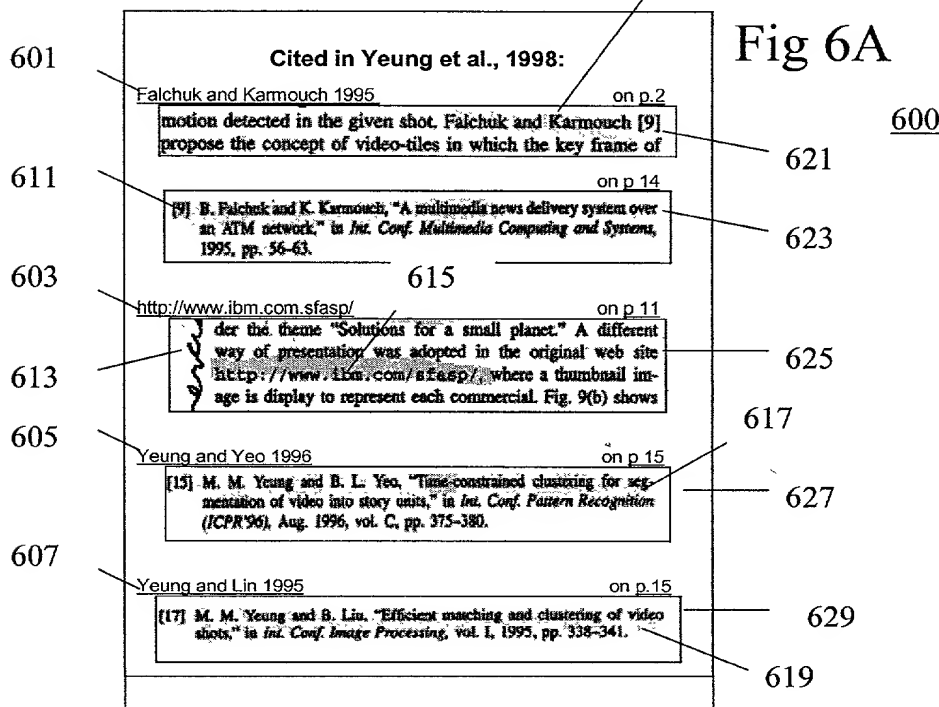


FIG. 5

609

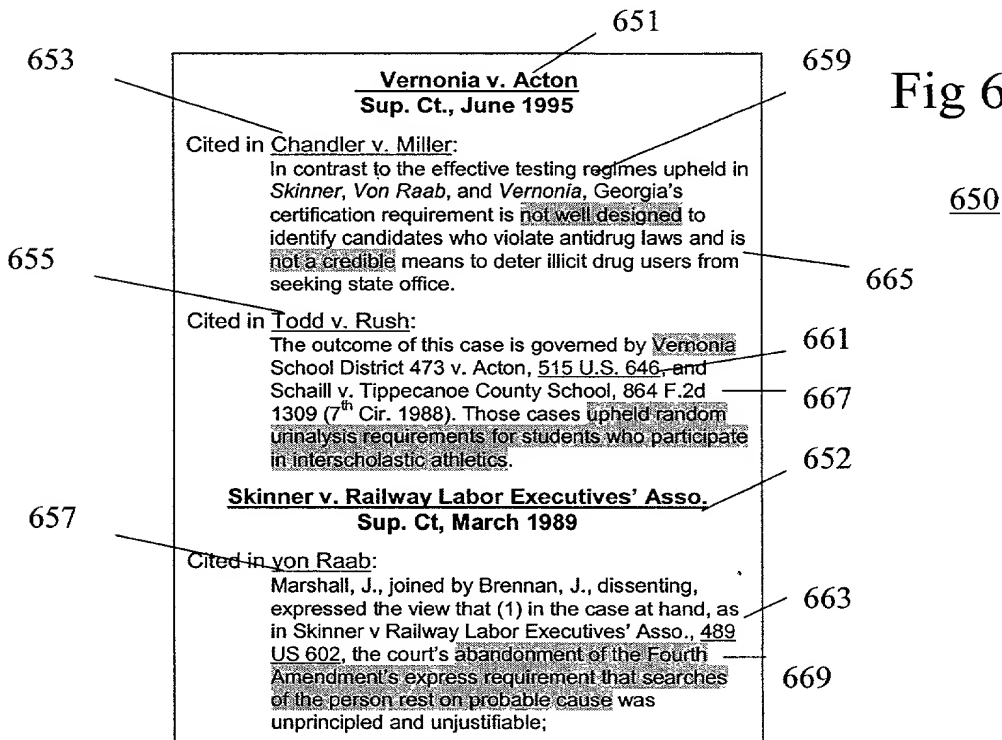
Fig 6A



651

659

Fig 6B



406

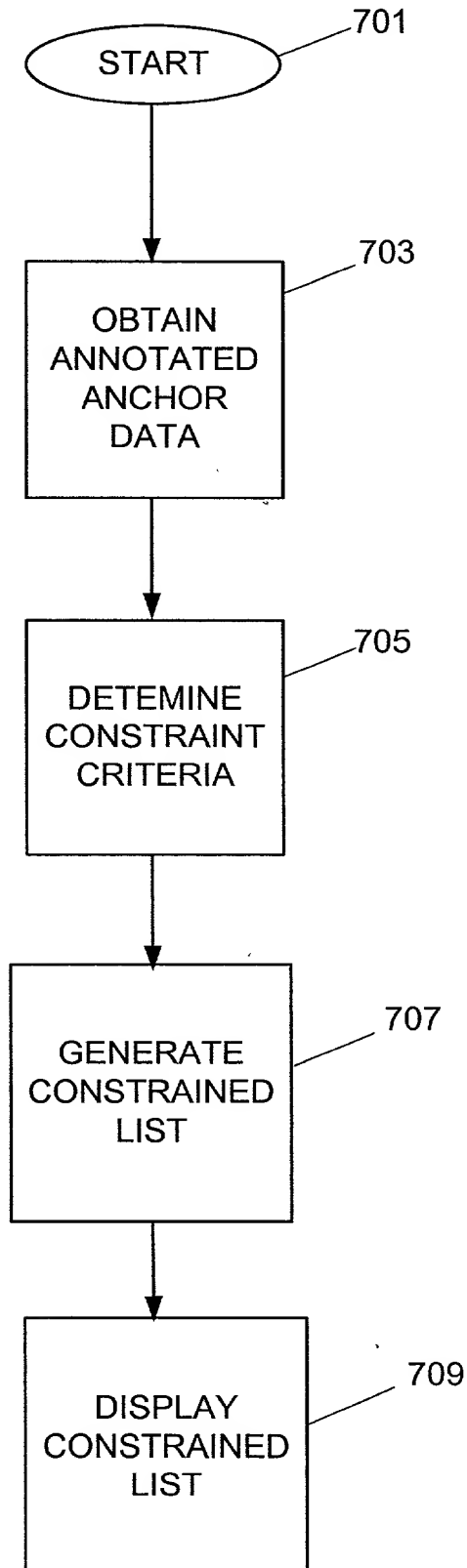


FIG. 7

408

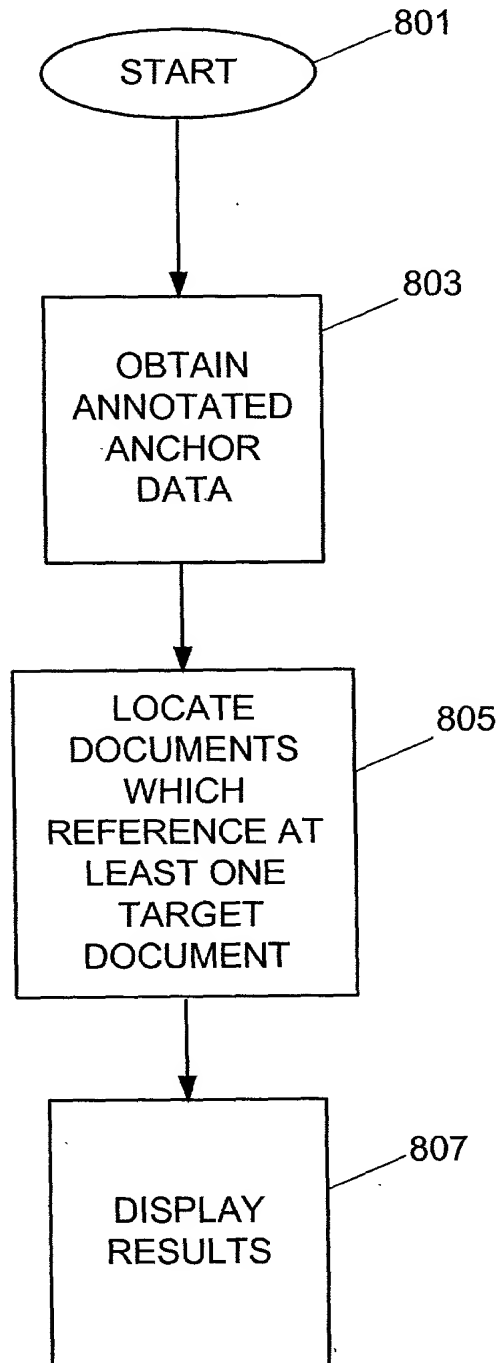
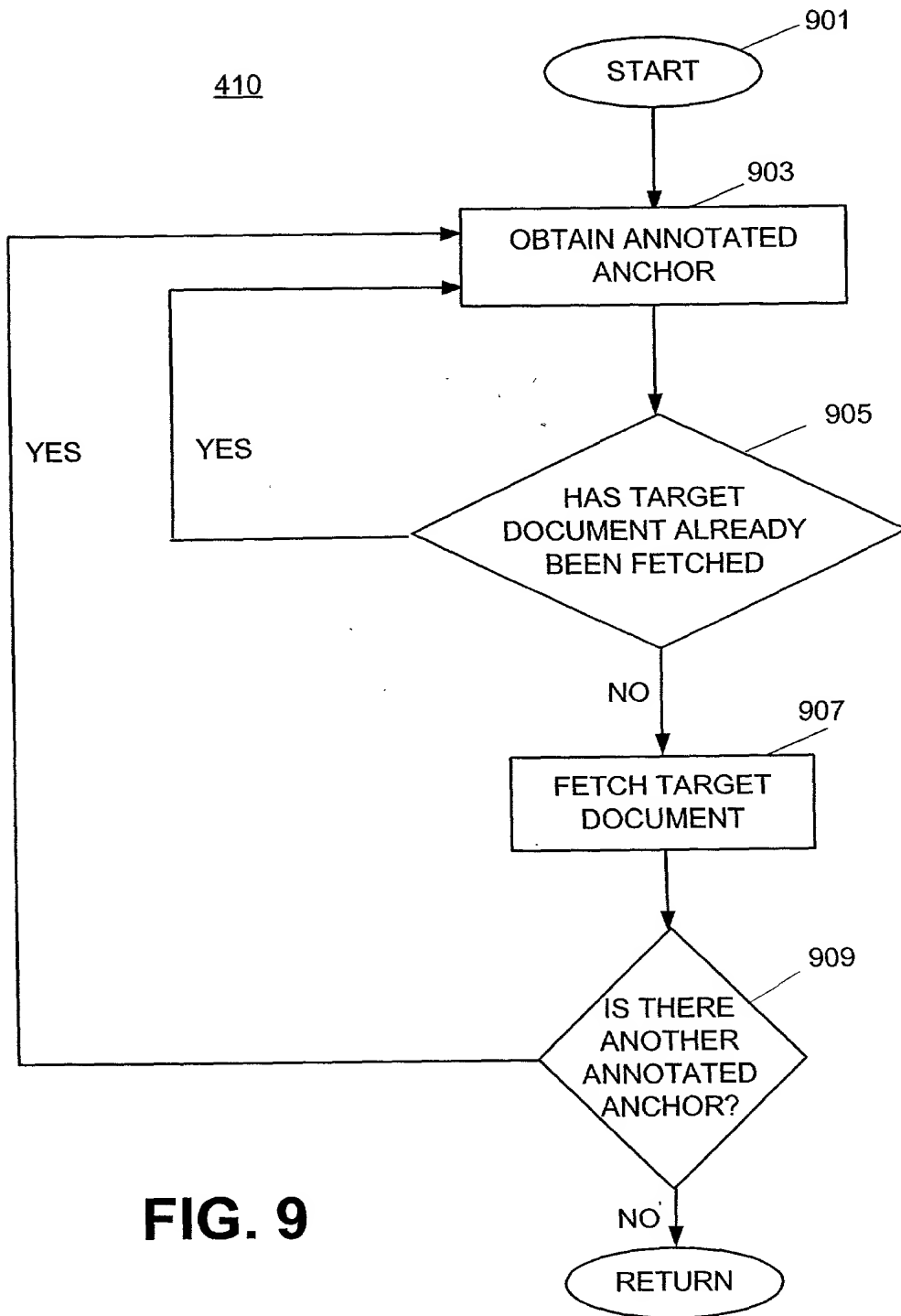


FIG. 8



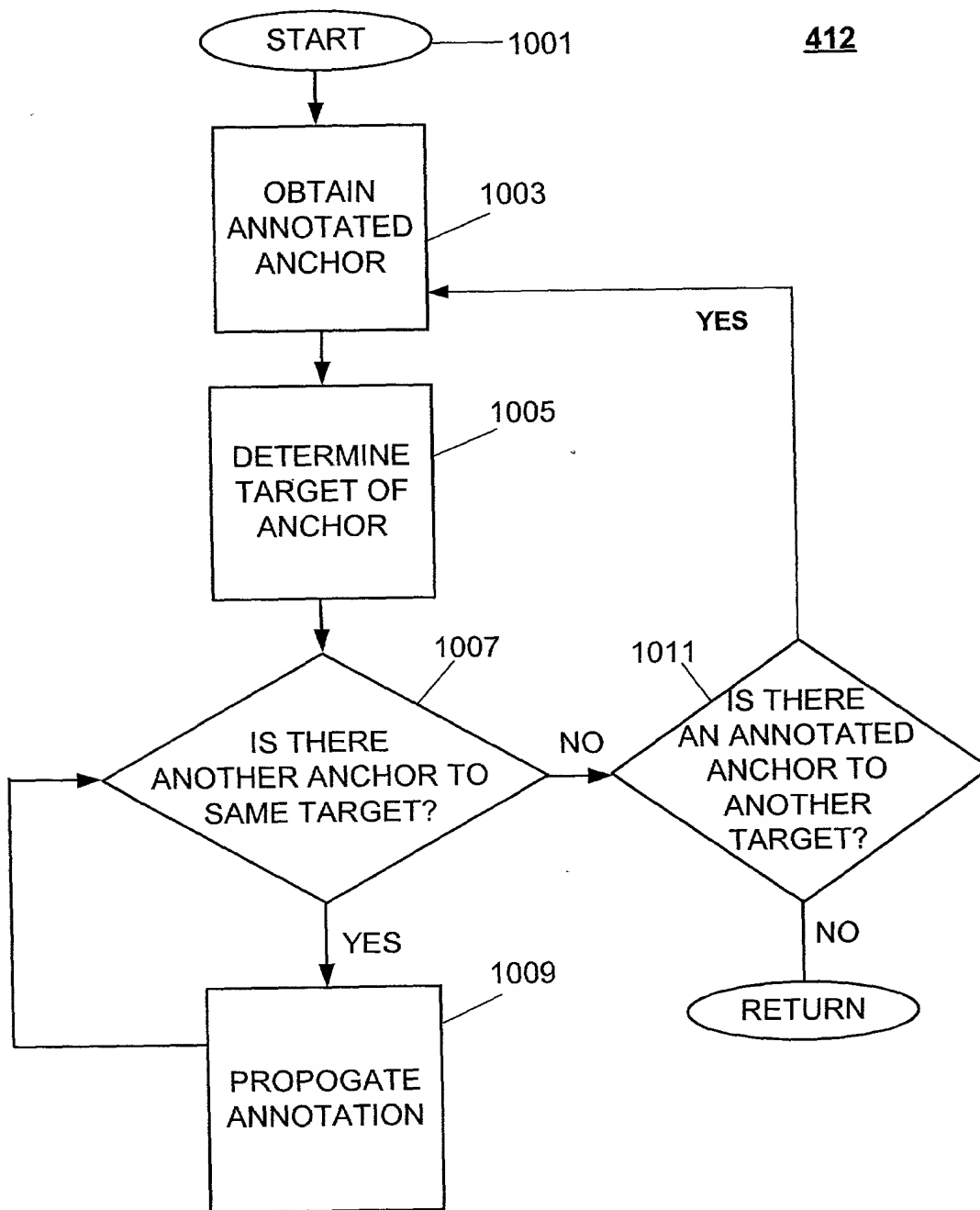


FIG. 10

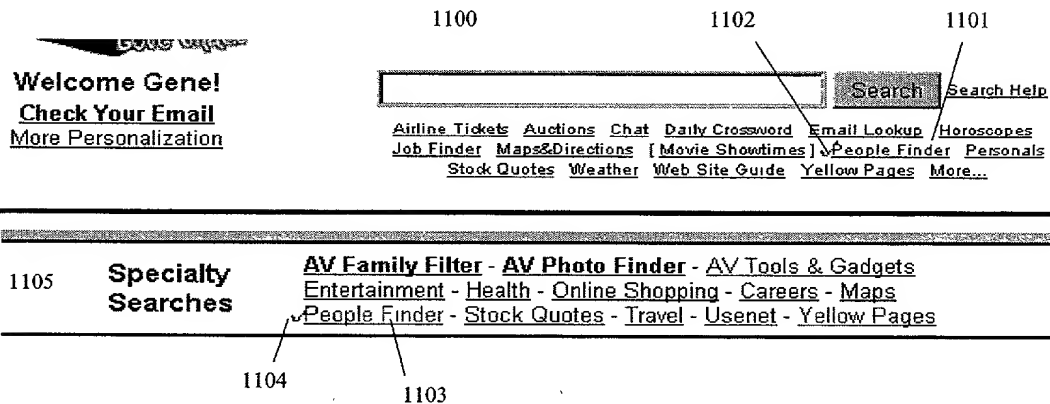


FIG. 11

418

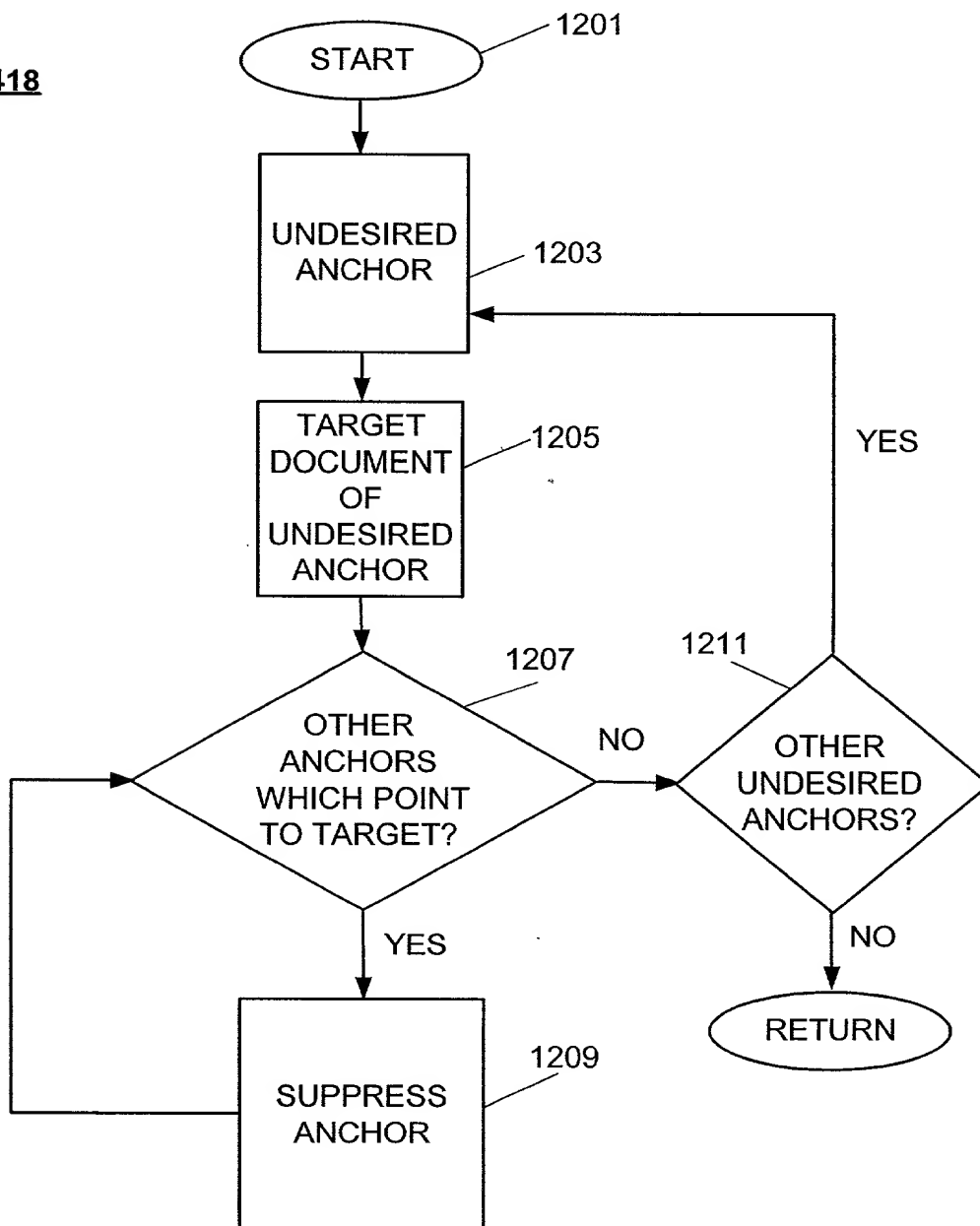


FIG. 12

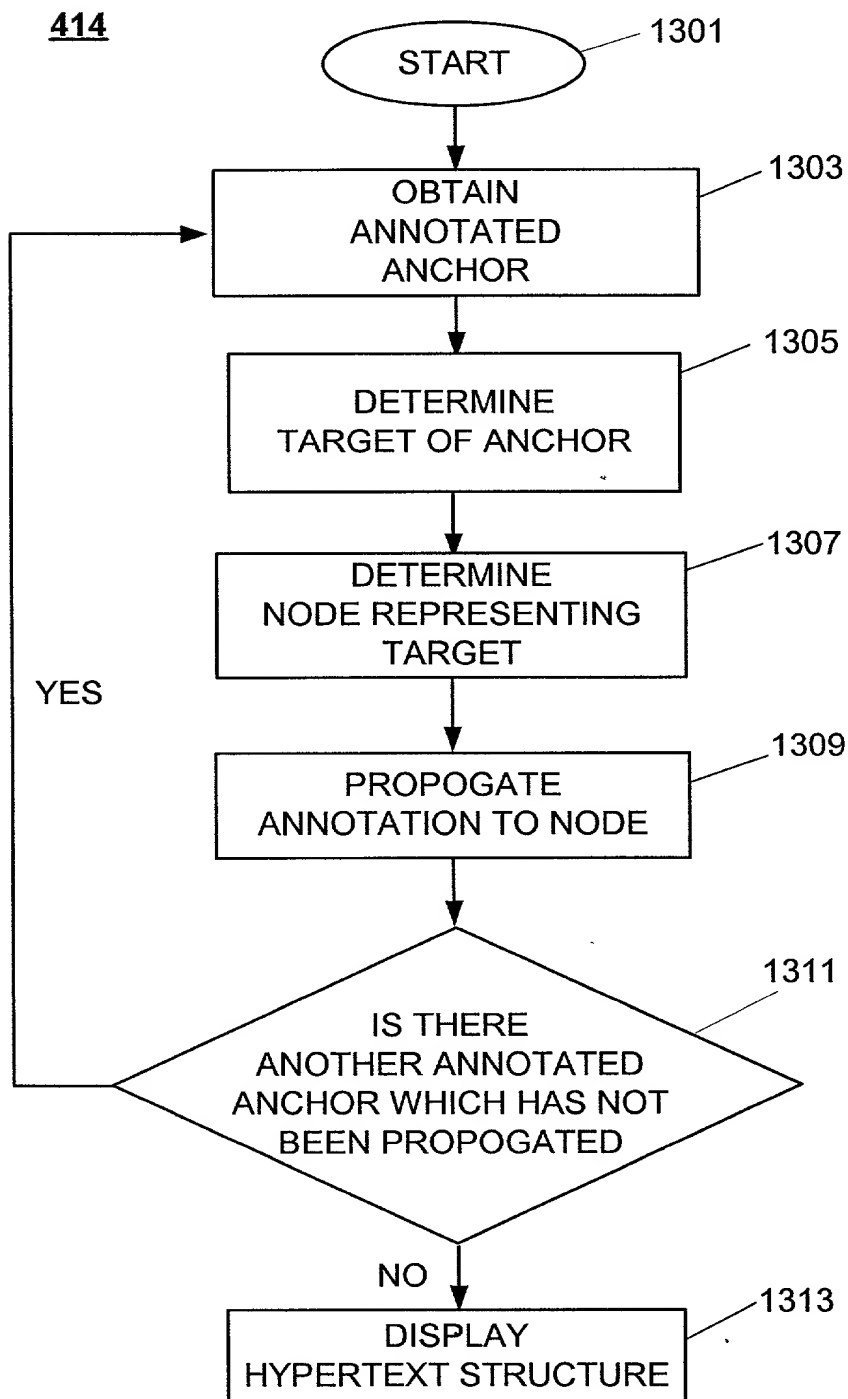


FIG. 13

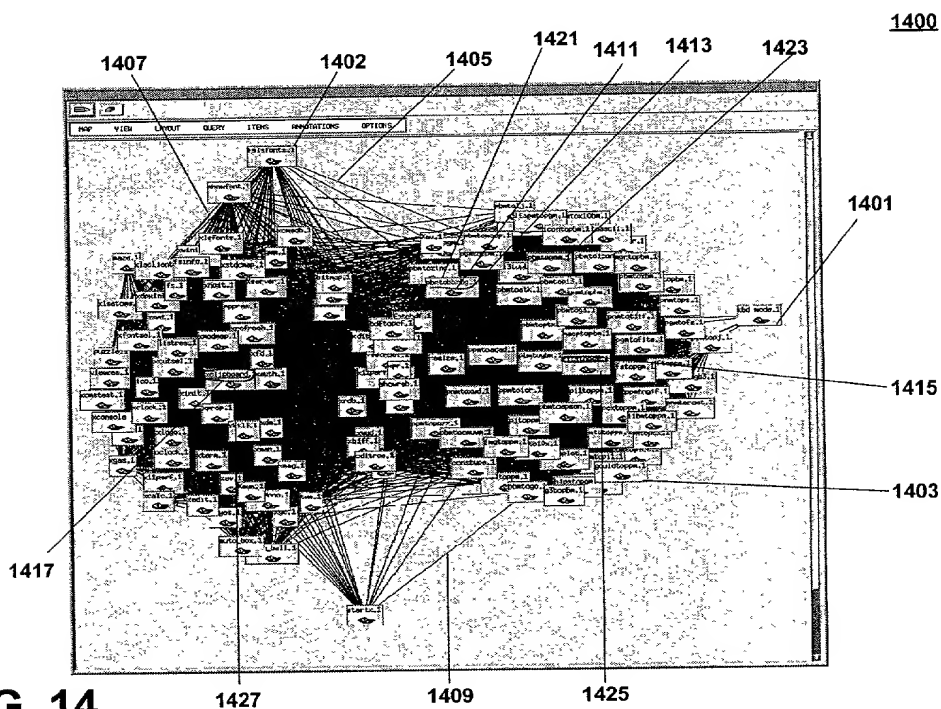
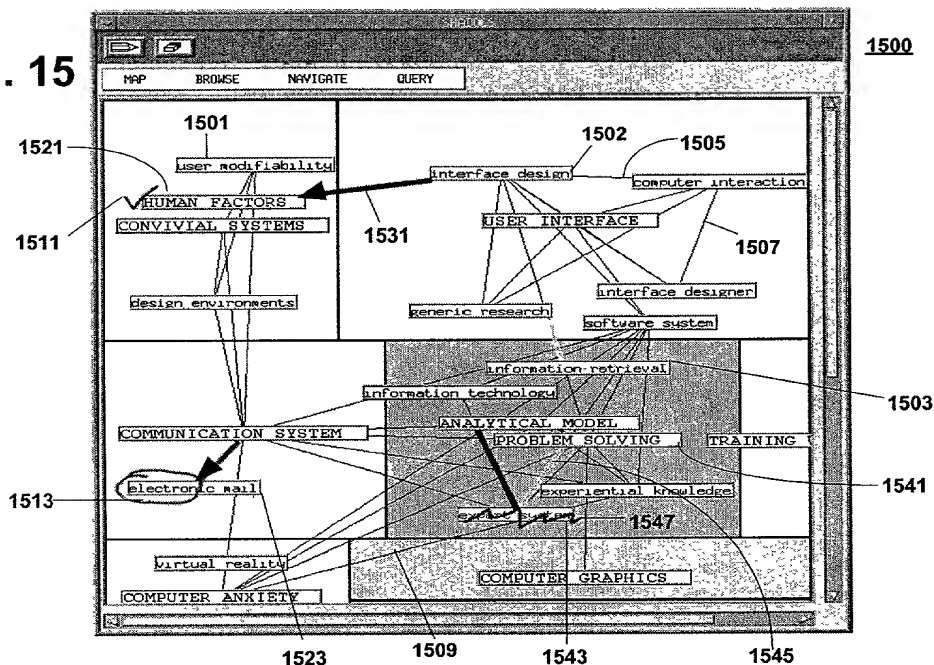


FIG. 15



416

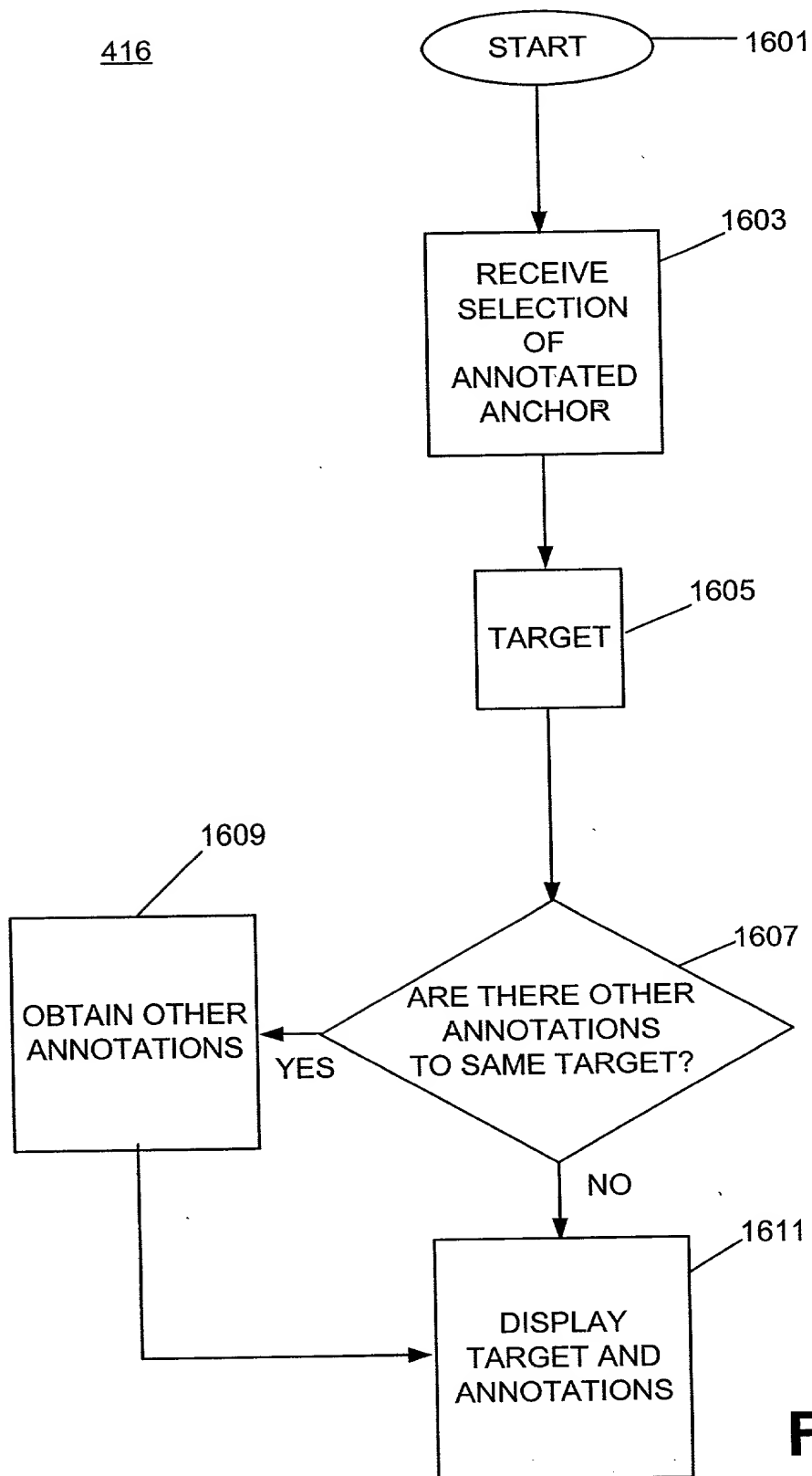


FIG. 16

09544817 083101

a shot with irregular lines as an indication of the camera motion detected in the given shot. Falchuk and Karmouch [9] propose the concept of video-tiles in which the key frame of

[9] B. Falchuk and K. Karmouch, "A multimedia news delivery system over an ATM network," in *Proc. Conf. Multimedia Computing and Systems*, 1993, pp. 56-63.

The second problem affects the reader more than the author. Very large collections of texts can be expected to contain vast numbers of potentially useful links. Unfortunately, most links would not be useful most of the time, and their sheer number can overwhelm even the most careful user interface design. Thus alternatives to semi-automatic link generation must be found.

1.2 Information retrieval and hypertext

The link creation problems described above suggest that more traditional information retrieval techniques must be used to accommodate large collections of nodes. Information retrieval methods have been applied to hypertext databases in several ways: link information has been used to inform retrieval algorithms (e.g., [8], [14], [19], [13], [36]), term co-location information has been used to suggest links to human authors (e.g., [2], [4], [32]), and queries have been used to retrieve hypertext nodes (e.g., [9], [12], [6], [5]). Relevance feedback has also been used to guide retrieval and to infer links among documents (e.g., [3]). Most systems that use queries as navigational aids use them to identify relevant neighborhoods in the hypertext, and then rely on manually-created links to support further navigation.

IR techniques have been used to segment long articles into shorter, more focused nodes (e.g., [34], [22]). Similarity among passages has been used to create links between specific nodes. This work, however, has focused on text segmentation techniques rather than on the hypertext interface. Although it is clear that such approaches are promising, little evidence has been published to date regarding their integration into interactive hypertext systems and about the effectiveness of such techniques in support of interactive browsing.

SuperBook, one of the more successful query-mediated browsing systems, used keyword queries instead of static hypertext links as a navigation mechanism [31]. Information was presented to the user in several windows, including table of contents (TOC), query, and text viewed. Users could use the TOC hierarchy to arrive at the desired section, or they could type in queries (or select keywords in the text). Search results were used to annotate the TOC to indicate relevant passages. Thus the system achieved hypertext-like browsing by combining TOC-based navigation with full-text search.

SuperBook's reliance on the table of contents to organize the browsing session limits it to providing access to highly-structured documents. Although an extension to SuperBook that works across documents

has been demonstrated [28], it still relies on hierarchical structure of each document to support local navigation. SuperBook has been shown to be an effective interface for IR tasks when browsing structured collections [12], but alternatives to the book metaphor must be found to support browsing through loosely-structured hypertext collections.

One such alternative — the newspaper metaphor — is discussed in the following section, and VOIR, a prototype that implements it, is described. Some experimental results from an evaluation of VOIR are presented, and the paper concludes with a discussion of possible extensions and applications of this query-mediated hypertext interfaces.

2 VOIR

This section describes VOIR (Visualization of Information Retrieval), a prototype newspaper-based dynamic hypertext interface. The section first introduces the newspaper metaphor and discusses its implementation in VOIR. A description of VOIR's linking interface follows, and the discussion concludes with an overview of VOIR's visualization features.

2.1 The newspaper metaphor

Newspapers such as the Wall Street Journal are designed to present a variety of different, loosely-related articles in a manner that supports browsing and selective reading. The front page of each newspaper section provides an overview of the contents. It presents summaries of articles, with references to other pages where additional details are discussed. The layout of each broadsheet provides cues to the relative importance of articles: important articles are usually placed near the top of the page, and more column space is allocated to them. These layout features serve to alert the reader to potentially useful information, and to structure interaction with a text that does not possess an overall narrative.¹

These features of a newspaper make it an appropriate vehicle for displaying hypertext information [18]. Users can capitalize on their familiarity with newspapers to browse hypertext collections. In addition to providing similarity-based structure, the newspaper metaphor can support the notion of landmark nodes [29] and hypertext links. The front page of a newspaper serves as a landmark around which semantically-related articles are organized. Articles split among several pages are connected with links. Overviews of contents are quite common. This

¹Each article, of course, has an internal structure. The newspaper merely serves to bind those largely-independent narratives together.

FIG. 17